HENDERSON HOSPITAL DEMOCRATIC THERAPEUTIC COMMUNITY: OUTCOME STUDIES AND METHODOLOGICAL ISSUES

Fiona Warren and Kingsley Norton

INTRODUCTION

History of Henderson Hospital

Henderson Hospital is a democratic therapeutic community, based in South London. It was set up in 1947 by Dr. Maxwell Jones, one of the pioneers of therapeutic community developments within the UK. The aptly named "Industrial Neurosis Unit" was started on an experimental basis with the implicit assumption that its utility required evaluation. In establishing the unit, the then Ministry of Labour's aim was to help resettle "the industrial misfit" following World War II (Whiteley 1980). The idea behind the experiment was simple: if during the war, a useful place had been found for society's social and psychological casualties, so it should also be possible in peacetime for them to be helped to find resources within themselves to lead more useful and fulfilling lives.

Over time, the patient group, treatment methods and aims and objectives of the Unit changed, as did its name. It became the Social Rehabilitation Unit in 1954, in acknowledgement that the inability of many to settle in a job was due more to their lack of social skills than industrial or educational expertise. Patients were increasingly referred via the Courts, rehabilitation offices or social work agencies to which they had drifted and become labelled as social inadequates. By the mid-fifties the Unit was recognised as a centre specialising in therapeutic community ideology and the treatment of "psychopaths" (Whiteley 1980). In 1959, the Unit was again renamed, Henderson Hospital, when it became an independent of its large parent hospital – Belmont Hospital. Its naming honoured Professor D.K. Henderson, author of "Psychopathic States" (Henderson 1939).

Over the ensuing forty years, the Henderson Hospital approach has stabilised, maintaining an equilibrium which continues to address the therapeutic needs of its severely personality disordered clientele, despite periodic threats to its survival resulting from its highly specialist (and unusual) status. The treatment has been described in increasingly greater detail, enhancing the chances that the treatment approach could be replicated and the generalisability of it's outcome findings evaluated (Whiteley 1986; Norton 1992a; Norton 1992b; Norton 1995; Norton 1996) and more recently, the hospital has collaborated with two other UK Mental Health NHS trusts (South Birmingham and Mental Health Services of Salford) in an attempt to replicate the model.

Henderson Hospital today

The aim of much of the integrated socio- and psycho-therapy programme can be summarised as converting people who "act", rather than thinking or feeling, into individuals who are able to sustain

and recognise their feelings and begin to articulate their psychological needs (Norton and Dolan 1995a). Maintaining this approach throughout 24 hours per day is problematic, entailing as it does staff working in close collaboration with the patients themselves (Norton and Hinshelwood 1996). Achieving this overall aim necessarily involves staff in an ongoing set of transference-countertransference relationships with patients, which spans (potentially) 24 hours of the day (Whiteley 1986). This psychodynamic situation requires an elaborate set of supportive and supervisory structures for staff in order for them to sustain, reflect on, and process their countertransference reactions (Rosenbluth 1991). With this kind of client group such activity is extremely important in minimising the destructive effects of any inter-staff splitting and promoting a therapeutic culture (Main 1957; Norton and Hinshelwood 1996).

Early literature on the TC suggested various definitions which may still be seen to underlie the approach today: 'a community with the immediate aim of full participation of all its members in its daily life with the eventual aim of re-socialisation of the neurotic individual for life in ordinary society' (Main 1946). For Maxwell Jones, the TC implied that: "the responsibility for treatment is not confined to the trained medical staff but is also the concern of other community members, i.e. the patients' (Jones 1952). Perhaps Main captured the democratic TC's essence with this definition:

This last definition also encapsulates the high degree of complexity within the treatment model and implies some of the difficulties with both its accurate description, or "manualization" for example, and its adequate evaluation (Norton 1992a).

The Past: Retrospective Outcome Studies

Social criteria of successful treatment: clinical improvement, employment, reconviction and readmission to psychiatric facilities

Clinical improvement

The Rapoport team, brought together to study the Belmont Social Rehabilitation Unit, intended to evaluate the treatment programme for which critics of Maxwell Jones and his method were calling (Rapoport 1960). However, there was controversy over the diagnostic use of the term "psychopath" and, partly because of this, the study changed direction, becoming more of an in-depth exploration of the workings of the community than an outcome study (Whiteley 1980). The overall study period spanned four years (1953 to 1957) (Rapoport 1960; Manning and Rapoport 1976).

A small-scale outcome study was incorporated, however, and the findings suggested that the democratic therapeutic community method was not universally applicable, that some selection

process for patients was necessary and that the intensive social and interpersonal pressures could damage those with weak ego structures. Rapoport had also pointed out a conflict of aims between those of the therapeutic staff and those of the workshop instructors who were aiming for "rehabilitation". The findings were largely viewed negatively by the Unit's staff and Maxwell Jones himself left the Unit in 1959 (Whiteley 1980).

For the outcome study Rapoport's team personally interviewed and classified 64 patients, one year after discharge, according to whether they were 'improved', 'same' or 'worse' compared to when they entered the Unit. Forty-one per cent were considered improved, eighteen per cent unchanged and thirty-one per cent worse, on the stated criteria. Improvement was associated with longer duration in treatment, fifty-two per cent for those staying more than 200 days were judged improved. No objective measures were used nor were any comparison samples studied.

Reoffending, employment and responsibility

In a later outcome study, a postal survey of the probation officers of eighty-six consecutive male discharges, on probation or borstal licences, sixty-two per cent of those traced were free of further convictions, up to twenty-two months post-discharge (Tuxford 1961). The response rate of the study was eighty-four per cent. Assessment was made using a four-point scale completed by the probation officer: from 1 (-increased sense of responsibility, employment and no further offending) to 4 (-further offending, unemployment and lack of responsibility). Twenty-four per cent fell into the first category, thirty-one per cent into the second, twenty-eight per cent were in the poor outcome category and seventeen per cent considered as complete treatment failures (Dolan and Coid 1993). Overall this represented a fifty-five per cent success rate, on the assessment criteria.

A similar rate of success was found in a follow-up study of discharged men (number unknown) who were assessed at 9 months post-discharge (Taylor 1963). Twenty-two per cent had found their own employment and a further forty-seven per cent had been placed in employment with professional support. The latter group was followed up for a further 9 months and sixty per cent of these were still in work with a satisfactory report from their employers.

Forty-five patients had no further psychiatric admission or conviction at 2- year follow-up. Of men with previous convictions, forty per cent remained free of conviction. Of those with previous admissions to a psychiatric hospital, almost sixty per cent remained out of hospital over the two-

year period. Good prognostic factors were: some level of school achievement; ever sustaining employment for more than two years; higher social class occupation; ever having been married and a history of affective disorder. A negative outcome was associated with: having previous convictions; a prior prison sentence(s); a probation order at referral or admission; current court proceedings; and institutionalisation before the age of 15.

A further outcome study was undertaken to develop a prediction equation for successful outcome (Copas and Whiteley 1976). Two cohorts of male patients were studied. One cohort of 104, at two years, showed forty-two per cent as having no further convictions or re-admissions, while the cohort of eighty-seven revealed a slightly higher figure of forty-seven per cent successful outcome on these criteria. At 5 years follow-up of the 104, one third had no reconviction or readmission. A further eleven per cent had only minor 'relapse' in the first year of follow-up, remaining free of conviction or relapse in the succeeding four years. It could be considered therefore that forty-five per cent of the total had a good outcome.

None of these early outcome studies utilized any comparison groups making it impossible to conclude that improvements were attributable to Henderson Hospital's democratic therapeutic community. However, a 5-year follow-up study of 194 (male and female) patients was carried out which also reported on fifty-one patients referred to Henderson but not admitted (Copas, O'Brien et al. 1984). Similar criteria of success were utilised. At 3 years, forty-one per cent of the treated sample was improved compared to twenty-three per cent of the non-admitted. At 5 years, the relative proportions were thirty-six and nineteen per cent. There was no significant gender difference. Further analysis of these findings showed a positive association between success rate and length of stay in treatment. At 3 years' follow-up, sixty-two per cent of those who stayed for 6 months and seventy-one per cent of those staying for 9 months, were improved. At 5 years, the relevant proportions were fifty-seven per cent and sixty-five per cent for 6 and 9 months' stays, compared with nineteen per cent of those not admitted (i.e. having treatment as usual elsewhere).

Psychological criteria: neurotic symptomatology

Reconviction and readmission to psychiatric facilities are often used as outcomes but are only indirect measures of psychological health. These earlier studies of outcome (measured mainly in terms of further conviction or psychiatric hospitalisation), were later complemented by the undertaking of a study of psychological morbidity (Dolan et al. 1992), as measured by the Symptom Check-list (SCL-90-R; Derogatis, Rickels et al. 1976). Sixty-two subjects were followed up at an

average of 8 months post-discharge and findings revealed a highly significant improvement on the Global Symptom Index (i.e. total score on SCL-90-R) post-discharge. Again, this was not a controlled study and the numbers were relatively small, although representing a typical figure for the number of patients treated in a given year.

The SCL-90-R is a self-rated instrument. However, data were subjected to a rigorous statistical analysis, addressing both the issue of reliability and also the importance of the observed clinical change (Jacobson and Truax 1991). With this method, fifty-five per cent of the group had improved reliably (i.e. had moved two standard deviations in relation to their baseline score). In thirty-two per cent, the change was clinically significant (i.e. subjects no longer scored in the pathological range but had scores within the normal range for the measure, defined by reference to normative data published on the instrument). Only 6.5 per cent had deteriorated reliably.

In this sample, length of stay was not significantly related to outcome, although those who stayed longer than nine months to show greater or beneficial change than those who stayed less (mean improvement .73, s.d. .84 for the longer stay vs .58, .86, for he medium stay and .61, 1.1 for the short stay groups). However, gender was significantly related to length of stay, more women staying longer than nine months (23/44, 52% vs 12/51, 29% chi squared 4.24 p=.04).

The present: Prospective Outcome Study

In 1990, a large prospective outcome study was launched by the then single-handed researcher, Dr. Bridget Dolan, who was at the time based solely in the hospital with no other formal academic links. Given the slender resources, this was an ambitious project. The study attempted to produce a psychological profile of all referrals during the study period. In addition it aimed to rate all patients admitted at 3 monthly intervals and at one and three years post-discharge; and those not admitted at one and three years post-referral. At this time there was no published large-scale prospective study in the field of personality disorder and the proposed study was breaking new ground in terms of improved research methodology.

Shortly after the study began, in 1991, major changes to the UK's National Health Service took place. In effect, service provision of the NHS was separated organisationally from its purchasing. As a consequence of the changes, local districts became more responsible for identifying the health needs of 'their' catchment area populations and using their funding allocation from central

Government to purchase these. Henderson's national catchment area meant that its funding depended on referrals other than from its local catchment area ('extra-contractual referrals'). Henderson's client group did not compete well with other extra-contractual referrals in the new 'market-place' (Dolan et al, 1994).

The numbers of applications for admission that attracted funding reduced, the financial viability of the hospital was threatened as, in the minds of purchasers, the treatment was not considered of proven worth but an expensive luxury which NHS could no longer afford. Paradoxically, two research benefits emerged from this otherwise negative scenario. First, a study of cost-offset was evoked in order to challenge the 'expensive' price tag (Menzies et al. 1993); (Dolan et al. 1996). Second, a group of referrals emerged who did not get admitted for treatment solely because the health authority refused to fund their ECR. This, it could be argued provided a comparison group closer to a randomly allocated control than the whole group of non-admitted referrals, for whom non-admission may have resulted from characteristics also relevant to treatment outcome (Dolan et al. 1997).

Cost as a measure of outcome

Two papers report the cost-offset of Henderson treatment. In the first study service usage data were collected retrospectively from May 1992 on 29 consecutive admissions and success rates of previously published outcome studies were used to calculate cost offset. In the second study actual service usage in the one-year post treatment was used to calculate the actual cost-offset for these admissions.

Data on mental health and forensic service usage in the one year prior to admission to Henderson derived from three sources: (1) case notes including information provided by the referrer; (2) the "Social History Form", a questionnaire completed by all admissions, concerning family, personal and clinical history; (3) subjects who were resident during the study period completed an additional form about the previous year's usage of services. Costs were calculated by obtaining figures of extra contractual referrals (ECR) tariffs for 1992/3 from the then four Thames Regional Health Authorities (RHA).

The daily tariff for Henderson Hospital was £111 compared with £153.20 for a general acute psychiatric bed and £173 for Close Supervision Units. The twenty-nine Henderson admissions had used a considerable amount of health and prison services in the previous year, the average estimated costs were £423,115 per year (mean cost per patient £14,590). In this first study, the cost of treatment at Henderson was offset by extrapolating from the 41% success rate in the studies reviewed above. This suggested that the treatment would pay for itself in four years (Menzies et al, 1993).

Subsequently in the second study, follow-up information had been obtained in the course of the prospective outcome study on actual service usage for 24 (83%) of the 29 residents in the original sample (Menzies et al, 1993). The average cost of services used by these twenty-four residents in the one year prior to admission was £13,966.

Information on service usage in the one year subsequent to discharge from Henderson was obtained from their referrer (in 17 cases) and/or their G.P. (in 14 cases) and from the client themselves (in 7 cases). Four subjects had further in-patient admissions, one of whom was re-admitted to the Henderson. Two people had outpatient assessments, twelve had outpatient treatment and one attended a day hospital. None of the residents spent time in prison or a secure psychiatric unit during the year. The average cost of services used was £1,308. This represents an annual saving post-discharge of £12,658.

These 24 residents were in treatment at Henderson Hospital for an average of 231 days (range 1 to 365 days) thus the actual cost of their treatment at Henderson was £25,641. Should the cost-offset continue at a similar rate for subsequent years then the cost of admission to Henderson would be re-couped in less than two years and represent savings thereafter.

Borderline symptoms as a measure of outcome

Reviewing outcome studies in the field of personality disorder (Norton and Dolan 1995c) note that many studies fail to assess the impact of treatment on aspects intrinsic to the personality disorder pathology itself, separately from those that are only associated or indirect phenomena. Indeed, there is a range of features associated with personality disorder, changes in which are erroneously equated with change in the personality disorder itself, such as reduction in axis 1 diagnosis symptomatology, or behavioural features such as criminal activity, self-mutilation or

suicidality. This prospective study aimed to assess changes in core personality disorder features. Comparison was made between those admitted and those not admitted for treatment. Consecutive referrals were mailed a self-report questionnaire pack on referral, including the Personality Disorder Questionnaire (PDQ-R) (Hyler et al. 1987), the Borderline Syndrome Index (BSI) (Conte et al. 1980), the Irritability, Depression and Anxiety Scale (IDA) (Snaith et al. 1978), and the Rosenberg Self-Esteem Scale (RSE) (Rosenberg 1965). A second follow-up assessment pack was set one year after referral (for the not-admitted) or discharge (for the admitted group). Up to three repeated mailings were used to maximise response rate.

The results of this study showed a significantly greater reduction in BSI scores in the treated than in the non-admitted group (Dolan, 1997). Changes in BSI scores were significantly positively correlated with length of treatment in the admitted group. Again, assessment of the reliability and clinical significance (Jacobson and Truax 1991) of changes in individual subjects was conducted in this study. These showed that the magnitude of this change was reliable and clinically significant in 42.9% of the admitted sample, compared with only 17.9% of the non-admitted sample (18.2% of the unfunded group).

Mood as a measure of outcome: Work in progress

Further results from the prospective outcome study described above are reported here. Given the positive effects shown on Borderline symptomatology, the outcome in terms of mood was also of interest. Depression, anxiety and inwardly and outwardly directed irritability were assessed using the IDA (Snaith et al. 1978) on the same cohort of patients described above. This instrument was chosen because it covers relevant symptoms theoretically related to personality disorder: anxiety and depression and, in addition, assesses the socially relevant tendencies to hurt the self or others. The scale is an 18 item scale with items scored from 0-4. The depression and anxiety subscales are scored between 0 and 15 and the irritability scales are scored between 0 and 12.

Participants

Response rates

The study sample is derived from the sample on which borderline symptoms were reported by Dolan, Warren et al. (1997). However the study period was extended therefore the sample and response rates will be summarised here. Consecutive referrals to the service between September 1990 and December 1994 were approached to participate in the study. There were 585 eligible referrals in the study period. Three hundred and eighty-four referrals (66%) completed a baseline assessment. Twelve of these participants (3%) were excluded from follow-up because they were

rereferred for treatment in the period between initial and follow-up assessment. One hundred and thirty five (36%) of the 372 eligible, responded at follow-up assessment. Seventy-five (56%) of the follow-up respondents were admitted and 60 (44%) not admitted.

Demographic characteristics

The mean age of the sample was 28 (range, 17-49, sd=6.8). Just under one half was female. Almost all the sample was white, single and unemployed. A large proportion had been previously convicted and had histories of drug and alcohol abuse. There were also high rates of previous suicide attempts, self-mutilation and overdosing reported by the referrers.

Personality disorder

This sample is a severely personality disordered group of people. The mean number of personality disorders for which each individual met criteria was seven and 95% of the sample met criteria for more than one personality disorder. Two cases did not score above threshold for any personality disorder and one case had missing data. The mean PDQ-R total score was 58 (see Table 1).

Table 1 Personality disorder symptoms: breadth and number of diagnoses

	Participants (n=134)
PDQ-R total score	
Mean (range), sd	**57.49 (17-86), 13.00**
Number of diagnoses[1]	
Mean (range), sd	**6.80 (0-12), 2.94**
More than one diagnosis	
Number of participants (%)	**127 (94.8)**

Eighty-four per cent of these participants met criteria for Borderline Personality Disorder.

Table 2 shows the prevalence of personality disorder diagnoses. Following Borderline, the most common diagnoses are Paranoid and Schizotypal, and Avoidant. The two research categories (self-defeating and sadistic) are the least prevalent.

[1] The research categories are included so that this is the number of diagnoses out of a possible 13

Table 2 Personality disorder sub-category diagnoses

Sub-category diagnoses	% Admitted (n=74)	% Not-admitted (n=60)
Paranoid	**77**	**65**
Schizoid	**55**	**33***
Schizotypal	**74**	**62**
Antisocial	**58**	**67**
Borderline	**82**	**85**
Histrionic	**55**	**67**
Narcissistic	**39**	**40**
Avoidant	**69**	**62**
Obsessive	**47**	**38**
Dependent	**62**	**52**
Passive-aggressive	**47**	**33**
Self-defeating	**14**	**22**
Sadistic	**22**	**27**

** significant at p<.05*

The most prevalent diagnoses were spread across the putative clusters into which personality disorder diagnoses are grouped in DSM-III-R & DSM-IV. Each cluster contained at least one score for over 80% of the group. Over 70% of the participants met criteria for personality disorders in all three clusters. There was a slightly higher proportion of Schizoid personality disorder in the admitted than the non-admitted group.

Representativeness

Given the naturalistic nature of this study and the attrition of participants over time, tests were conducted to establish the representativeness of the sample. Group differences at baseline suggested that outcome analyses should take sex and the presence of schizoid personality disorder.

Length of stay

The average length of stay of the admitted participants in this comparison was 201 days (6.7 months). The minimum length of stay was 2 days and the maximum 396.

Outcomes

The results of repeated measures analysis of variance are summarised in Table 3. The table shows the baseline and follow-up mean score for each group and the interaction effect. The BSI results have been included here to provide continuity with the previously published results on this measure (Dolan et al. 1997). A highly significant interaction effect is shown.

Table 3 Summary of Outcomes

		Baseline	*One year*	F	P
Borderline Symptoms	Admitted	34.89 (9.39)	22.03 (15.29)	10.85	.001
	Not admitted	32.98 (11.55)	28.26 (12.90)		
Anxiety	Admitted	10.01 (2.73)	7.61 (4.04)	11.98	.001
	Not admitted	8.67 (3.07)	8.56 (3.44)		
Depression	Admitted	8.73 (3.09)	6.70 (3.98)	4.61	.034
	Not admitted	7.68 (3.78)	7.14 (3.13)		
Inward Irritability	Admitted	8.48 (3.28)	6.25 (3.73)	7.78	.006
	Not admitted	6.80 (3.53)	7.14 (3.13)		
Outward Irritability	Admitted	6.53 (2.78)	5.77 (3.16)	5.09	.026
	Not admitted	6.33 (2.87)	6.76 (3.02)		

The results show highly significant interactions for anxiety, depression, inwardly and outwardly directed irritability. The mean scores show that in each case this interaction is a result of greater improvement in the admitted group between baseline and one year follow-up. There is a suggestion of improvement in the non-admitted group for Borderline symptoms, anxiety and depression, but of a deterioration in inward and outwardly directed irritability.

The group differences at baseline on anxiety and inward irritability were significant with higher (more pathological) scores in the admitted group. Adjusting for this produced a non-significant interaction for inwardly directed irritability. There was a main effect of time, significant at the $p<.05$ level, however, suggesting that both groups were showing some improvement on this measure.

Confounding variables

When sex, and Schizoid Personality Disorder were entered as factors into the individual analyses of variance, no interactions or main effects of these variables were found. For outwardly directed irritability, there was no interaction of schizoid personality disorder with either time or group.

Relationship of outcomes to length of stay in treatment - Association of length of stay with follow-up score

Length of stay was negatively correlated with all follow-up scores: the longer a resident stayed in treatment the lower their follow-up score. These negative correlations were significant to $p<.05$ level with follow-up scores on borderline symptoms, anxiety, depression. The negative correlations with inwardly and outwardly directed irritability were not significant.

Association of length of stay with change

(Dolan et al. 1997) found a significant correlation between change in borderline symptoms and length of stay in treatment. These analyses also show a significant correlation between length of stay in treatment and degree of change between baseline and one-year post treatment follow-up for depression and anxiety. However, there was no signifcant correlation for inwardly and outwardly directed irritability.

Comparison of change for short-stay and long-stay participants

The admitted patients were therefore divided into long and short stay groups. Short-stayers were those who stayed less than three months and long-stayers those who stayed nine months or more in treatment. The short stay group stayed in treatment a mean 34 days (sd 29.9, range 2-91 days). The long stay group remained in treatment a mean of 343 days (sd 32.2, range 277-396 days).

T-tests revealed significant differences for the short and long stay groups in change scores for borderline symptoms, anxiety, depression, and inwardly directed irritability. All of changes are improvements for the admitted group. Differences in change scores were non-significant for outwardly directed irritability.

Earlier studies (Dolan, 1992; Copas et al. 1984) had found a significant gender difference in length of stay. This was not evident in this study.

Summary of work in progress

These results augment the existing evidence of positive treatment outcomes. Improvements in core personality disorder pathology previously shown (Dolan et al. 1997) are supported by the improvements in mood symptomatology shown here. Treatment effects seem to be shown in terms

of reductions in anxiety, depression and outwardly directed irritability, although reduction in inwardly directed irritability would seem to be a weaker effect. It is of interest that there were significant differences between the admitted and non-admitted referrals in terms of anxiety and inwardly directed irritability on which the admitted group scored more highly. This may suggest a selection effect, which should be explored by future research.

Comment

This prospective outcome study also suffers from some of the methodological shortcomings levelled at previous studies, above. The use of self-report measures limits the validity of the findings, although some reassurance can be gained from the consistency of effect across multiple self-report measures. The proportion of those about whom we have data at outcome is only a small proportion of the eligible sample in both groups. However, the response rate in this study is not atypical of a PD sample of patients. This limits the generalisability of the findings to PD patients in general. In addition, the follow-up interval differed between the treated and non-treated samples so that the non-treated sample was followed up earlier than those treated. Alternative study designs, which attempted to match the timing of a non-treated follow-up with a treated follow-up, would only have been possible in theory, since the time spent in treatment for any individual patient could not be known in advance! However, it is also highly unlikely that time alone accounts for such a magnitude of difference between the admitted and non-admitted groups when spontaneous remission in this client group is widely acknowledged to be rare. Further analysis of results using the data collected during treatment may help to substantiate this. The non-admitted comparison group cannot truly be labelled "non-treated" since it is likely that they had at least some non-specialist treatment during the study period which could not be controlled for. The use of this comparison sample is also problematic because the reasons for non-selection or non-attendance may relate to a poorer prognosis at the point of entry to the study. The use of the non-funded group in the study on borderline symptoms, however, provides some reassurance. The non-random allocation to treatment or non-treatment, is the most problematic methodological limitation of the study because this allows systematic variation between groups. Some of the difficulties of applying randomisation in this context are identified in the discussion below.

Summary of outcome studies of Henderson Hospital

Table 4 shows a summary of the outcome studies reviewed above. These have shown improvements in those admitted for treatment using a range of approaches to outcome measurement including convictions, psychiatric service usage and a various kinds of psychological functioning.

The proportion improved seems to be consistently around 40% but this may improve to around 70% for those who stay in treatment for nine months or more. Treatment gains have been shown to be maintained up to five years post treatment. There is some evidence that a small proportion 3-35% deteriorate following treatment. Where comparison groups have been used, they have been shown to fare significantly worse than those admitted to the treatment although a small proportion of "untreated" controls also show improvement over time. Some of the earlier studies were limited in terms of comparison groups, measures used and follow-up periods although the methodological approaches can be seen to have evolved over time.